Services and Technology Overview

PASIG March 13, 2015

Evviva Weinraub, Services Manager
David Pcolar, Technical Manager
Services and Technology Overview

PASIG March 13, 2015

Evviva Weinraub, Services Manager
David Pcolar, Technical Manager
What is DPN?

To avoid the catastrophic loss of scholarship, we must build and sustain a diverse ecosystem that can ensure the survival of scholarship in digital form for future generations.

We envision a system that is scalable, sustainable, and complementary to existing collection and preservation efforts—the Digital Preservation Network (DPN or Deepen).

- James Hilton
What is DPN?

Building a Dark Archive (Darkive?) for the Academy
DPN Technical Partners

Ingesting/Administrative & Replicating Nodes
- Academic Preservation Trust (APTrust)
- Chronopolis
- DuraSpace
- Stanford Digital Repository (SDR)
- University of Texas Data Repository (UTDR)

Ingesting/Administrative Node Only
- HathiTrust
What Does DPN Do?

- Establishes a network of heterogeneous, interoperable, trustworthy, preservation-level dark archives (darkive)
- Replicates content across the network, to multiple nodes
- Enables restoration of preserved content to any node due to data loss, corruption, disaster, or other catastrophic event
Pathways to Ingestion

Your Content

Ingesting Node

DPN Bag
Welcome to the Darkive
(We’ve Got Cookies!)

DPN Bag

Transfer

DPN Ingest Node

Replicating Node
In case of catastrophic loss...
Content Restoration

Replicating Nodes

DPN Bag

Your Content
Architectural Overview

• Architectural Premise
  o Core capabilities founded on proven institutions and repositories

• Design Considerations
  o Distributed Nodes, loosely coupled
  o Standards and protocol-based integrations
  o Separate implementations
  o Distributed infrastructure

• Outcome is a safe preservation environment - DARKIVE!
Current State of Affairs

- Successful Pilot/Proof of concept run in late 2014
  - Baggit submission
  - Ingestion
  - Replication
- Hired Technical Manager in December
- Hired Services Manager in February
- Working on implementing Agile methodologies across development team
What Does Launch Look Like?

• Aiming for a July 1, 2015 soft launch
  • Depositor will be able to give us stuff and we can put it into storage
  • Some cursory reporting available to Depositor
  • Ability to replicate content from the Ingest Node to all Replicating Nodes
  • Ability to restore content on request
  • Replicated inventory of items in DPN
  • Fixity checks on ingest & replication
What’s NOT in Launch

...but we have time-lined out for this year

• Signed, reviewed, & accepted SLA’s with member institutions
• Business model released to community
• Production capacity content ingestion
• Clear idea of Node storage capacity & timelines for infrastructure onboarding
• Cost models with Ingesting/Administrative Nodes
• Ongoing fixity checks
• Reporting dashboard
We Know You Have Questions...

DPN is a 60+ member cooperative organization investing in long-term, scalable digital preservation for the academy.
http://www.dpn.org

Evviva Weinraub: evviva@dpn.org
Dave Pcolar: dave@dpn.org
What is DPN?

To avoid the catastrophic loss of scholarship, we must build and sustain a diverse ecosystem that can ensure the survival of scholarship in digital form for future generations.

We envision a system that is scalable, sustainable, and complementary to existing collection and preservation efforts—the Digital Preservation Network (DPN or Deepen).

- James Hilton
What is DPN?

Building a Dark Archive (Darkive?) for the Academy
DPN Technical Partners

Ingesting/Administrative & Replicating Nodes

• Academic Preservation Trust (APTrust)
• Chronopolis
• DuraSpace
• Stanford Digital Repository (SDR)
• University of Texas Data Repository (UTDR)

Ingesting/Administrative Node Only

• HathiTrust
What Does DPN Do?

• Establishes a network of heterogeneous, interoperable, trustworthy, preservation-level dark archives (darkive)

• Replicates content across the network, to multiple nodes

• Enables restoration of preserved content to any node due to data loss, corruption, disaster, or other catastrophic event
Pathways to Ingestion

Your Content

DPN Bag

Ingesting Node
Welcome to the Darkive
(We’ve Got Cookies!)

DPN Bag

Transfer

DPN Ingest Node

Replicating Node
In case of catastrophic loss...
Content Restoration

Replicating Nodes

DPN Bag

Your Content
Architectural Overview

• Architectural Premise
  o Core capabilities founded on proven institutions and repositories

• Design Considerations
  o Distributed Nodes, loosely coupled
  o Standards and protocol-based integrations
  o Separate implementations
  o Distributed infrastructure

• Outcome is a safe preservation environment - DARKIVE!
Current State of Affairs

- Successful Pilot/Proof of concept run in late 2014
  - Baggit submission
  - Ingestion
  - Replication
- Hired Technical Manager in December
- Hired Services Manager in February
- Working on implementing Agile methodologies across development team
What Does Launch Look Like?

- Aiming for a July 1, 2015 soft launch
- Depositor will be able to give us stuff and we can put it into storage
- Some cursory reporting available to Depositor
- Ability to replicate content from the Ingest Node to all Replicating Nodes
- Ability to restore content on request
- Replicated inventory of items in DPN
- Fixity checks on ingest & replication
What’s NOT in Launch

...but we have time-lined out for this year

• Signed, reviewed, & accepted SLA’s with member institutions
• Business model released to community
• Production capacity content ingestion
• Clear idea of Node storage capacity & timelines for infrastructure onboarding
• Cost models with Ingesting/Administrative Nodes
• Ongoing fixity checks
• Reporting dashboard
We Know You Have Questions...

DPN is a 60+ member cooperative organization investing in long-term, scalable digital preservation for the academy.

Arizona State University  Brigham Young University  Brown University  California Institute of Technology  Columbia University  Cornell University  Dartmouth College  Duke University  Emory University  Harvard University  Indiana University  Iowa State University  Johns Hopkins University  Kansas State University  Massachusetts Institute of Technology  Michigan State University  New York University  Northwestern University  North Carolina State University  Ohio State University  Pennsylvania State University  Princeton University  Purdue University  Rutgers University  Stanford University  Syracuse University  Texas A&M  Texas Tech University  Tufts University  Tulane University  University of Alabama  University of Arizona  University of Buffalo  University of California San Diego  University of Chicago  University of Florida  University of Illinois at Chicago  University of Illinois at Urbana-Champaign  University of Iowa  University of Kansas  University of Kentucky  University of Maryland  University of Miami  University of Michigan  University of Minnesota  University of Nebraska  University of New Mexico  University of North Carolina  University of Notre Dame  University of Tennessee  University of Texas  University of Utah  University of Virginia  University of Washington  University of Wisconsin  Utah State University  Vanderbilt University  Virginia Polytechnic Institute and State University  Yale University  Texas Digital Library  California Digital Library  John D. Evans Foundation  American Council on Education